UNCLASSIFIED

Defense Technical Information Center Compilation Part Notice

ADP010689

TITLE: Toward a Methodology for Evaluating the Impact of Situation Awareness on Unit Effectiveness of Dismounted Infantrymen

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Usability of Information in Battle Management Operations [1'Exploitation de l'information dans les operations de gestion du champ de bataille]

To order the complete compilation report, use: ADA389629

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, ect. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP010683 thru ADP010703

UNCLASSIFIED

Toward a Methodology for Evaluating the Impact of Situation Awareness on Unit Effectiveness of Dismounted Infantrymen

(April 2000)

Cynthia L. Blackwell
USA SBCCOM-NSC
ATTN: AMSSB-RSC-MA(N)
Natick, MA 01760
USA
and
Elizabeth S. Redden
U.S. Army Research Laboratory
ATTN: AMSRL-HR-MW
Ft. Benning, GA 31905-5400
USA

Introduction

The United States (US) Department of Defense initiated a program in 1997 called the Military Operations in Urban Terrain Advanced Concept Technology Demonstration (MOUT ACTD). MOUT ACTD is a joint US Army-Marine Corps. program led by the US Army Soldier and Biological Chemical Command. The MOUT ACTD's charter is to seek technologies that satisfy 32 jointly derived requirements specific to operations in 'built up' or urban areas. MOUT ACTD evaluates these candidate technologies for military utility and transitions the successful candidates to acquisition programs for further development and fielding.

One of the determinants of military utility that the MOUT ACTD program uses is situation awareness (SA) - specifically, the influence of SA on individual and force effectiveness as a result of the use of MOUT-related technologies. SA is defined here as the warrior's ability to quickly perceive and then discriminate between facets of the tactical environment, to accurately assess and reassess the where, when and why of that environment, to then know and understand the nature of the tactical situation and to extrapolate near term courses of action based on this understanding. This paper describes the process by which the MOUT ACTD program developed and implemented a method for determining the impact of SA on individual and force effectiveness.

Background

Behavioral scientists have investigated SA to a great degree in U.S. Navy, Air Force, and Army aviation communities. SA, as it relates to dismounted infantry operations, is an emerging area of study. The excellent research in this field conducted by the aviation community, in both individual and crew SA, has made it possible for behavioral scientists to formulate new approaches of investigation of SA for the dismounted war fighter and small infantry unit.

With the advent of advanced technologies and battlefield digitization, materiel developers are now required to provide connectivity between the dismounted infantryman (DI) and the digitized battlefield through novel communications, sensing and command, and control-enhancing technologies. For example, many U.S. soldiers may soon be provided wearable computers to aid individual and small unit command and control operations. The technologies incorporated in these computers include advanced sensors, communications, and navigation. These technologies have the potential to directly impact, both positively and negatively, the ability of the war fighter to perceive his or her environment and to understand his or her place in it. The war fighter must be able to interface with these technologies to best employ the capabilities they provide. Research needs to be done to first understand the impact these capabilities have on the warfighter's cognitive abilities and ultimately to interweave the output of these capability-enhancing technologies into the infantryman's decision-making processes that ultimately impact battlefield outcomes.

SA Measures of Effectiveness and Performance

There is a variety of factors that indicate how effectively a war fighter operates in a battlefield environment. How individual and small units employ new technologies is a function of this effectiveness. The operational effectiveness factors include casualty ratios, logistics resupply, combat support, lethality, and survivability. Another of these factors is SA.

Mission accomplishment is the critical operational factor in determining the relative military utility of a piece of technology. The basic premise of the MOUT ACTD SA effort is that the more situationally aware a force is, the more lethal, mobile, and survivable a force is.

Paper presented at the RTO HFM Symposium on "Usability of Information in Battle Management Operations", held in Oslo, Norway, 10-13 April 2000, and published in RTO MP-57.

Quantifying this was and is the ultimate goal of the MOUT ACTD SA effort.

In order to reach the goal of developing a methodology for quantifying the influence of SA on performance, it was necessary to find a way to determine its value in a field environment. The challenge was to define objective, field expedient, operationally based measures for determining the influence SA has on mission performance. To accomplish this, a panel of experts was convened at the US Army Research Institute, Ft. Benning Georgia, USA. The charter of this panel was to determine measures for SA that would quantify changes in combat effectiveness. The panel was composed of active duty and retired Army and Marine Corp. officers and enlisted personnel, training experts and behavioral scientists. The panel met on four separate occasions corresponding to scheduled experiments during which particular SA-impacting technologies would be used.

Goal-Directed Knowledge Elicitation Technique

A formal process was needed to solicit the required field-expedient measures. The panel used what came to be called the Goal-Directed Knowledge Elicitation Technique (GDKET) to accomplish this. This technique was developed as a way to solicit situation-specific mission needs from subject matter experts (SMEs). These mission needs were established to provide the means by which specific mission goals were reached and the knowledge of what must occur on the battlefield in order to reach those goals, which served as the foundation of the SA measures.

GDKET was developed based on the goal-directed task analysis described by Endsley in "Situation Awareness Information Requirements for En Route Air Traffic Control" (Endsley & Rogers, 1994). The general approach described by Endsley and Rogers is the same used here; however, the specifics of expert knowledge elicitation and requirements analysis differs in that GDKET uses a panel of experts to identify mission goals. GDKET also relies on the interplay of the experts during role-playing to generate requirements that form the basis of the SA measures of performance. The behavioral scientist serves as an observer and facilitator to cull the requirements from the panel's discussions in a non-invasive way. In GDKET, the experts assist in analyzing the requirements and devising the measures. Emphasis is placed on obtaining field expedient measures, and the GDKET approach allows these to be obtained.

In the GDKET mission vignettes were developed that incorporated individual and small unit tasks during which MOUT ACTD technologies could be employed. An example of one such mission vignette was clearing a building of enemy troops. The panel discussed the mission vignette in order to achieve a common understanding of the mission. Experts, in turn, were queried about their role in the mission (e.g., platoon leader, company commander, squad leader). After this had been established the vignette was "role-played," thereby soliciting the dynamic elements of the mission including tactics, techniques and procedures (TTPs) and standard operating procedures (SOP). All this was done independently of technologies employed during the mission.

The panel members were queried about their tasks as well as their intent at different levels of granularity with respect to the mission. For example, in the building clearing vignette, the squad leader was asked to describe his activities during the mission. He described his mission goal first (clear a floor in the building of enemy soldiers) then general mission tasks (e.g., providing status reports to his platoon leader). Based on the mission goal, sub-goals and supporting tasks were identified. An example of a sub-goal was to ensure that the squad has adequate supplies of ammunition and water during the mission. Supporting tasks included periodic querying of fire-team leaders about their ammunition status and passing that information to the platoon leader.

A list of operational requirements was generated and these formed the basis of the SA measures. The following is an example of some of the information requirements generated for a squad leader during a building clearing mission and the resulting measures:

Squad Leader must know:

- Location of platoon leader
- Location of other squads
- Ammunition, water and equipment status
- What rooms and floors have been cleared – how many are left
- Status of fire teams
- Room and floor layouts - blueprints
- Location of the enemy – what floor
- Location of non-combatants and animals
- Casualties
- Prisoner of war collection points
- Location of Platoon Sergeant for resupply
- Some of what the platoon leader knows

Measures:

- Measure the information actually reported against the expected information. Based on TTPs and SOPs the assumption is that communications within a unit will be 100% accurate. The reality is that some messages do not get through.
- Frequency of reporting. How often was a report presented and received versus when it was expected, based on TTPs and SOPs.
- Accuracy of report. Was the report received accurately; was it complete and did it contain the correct/intended information. This is a function of quality of information provided to/and presented by the sender

- Timeliness of report. Were the reports sent when required? This is based on TTPs and SOPs.
- Two-way communication. Quality and quantity of interaction.
- Reduced risk of fratricide.
  - -Unit-to-unit proximity
  - -Status and location of friendlies
  - -Number of wounded
- Number of correct decisions made. A function of advancing the mission, momentum - are the correct decisions made at the appropriate time?
- Speed and reliability of report.

The panel realized that some of the measures could be used near term to determine combat effectiveness, but some (e.g. reduced risk of fratricide) would require long-term experimentation to determine the real impact.

At the end of the GDKET exercise, the panel was made aware of the new capabilities that they would have available to them to use during the mission. For each mission vignette, the experts discussed the impact of these capabilities on their goal, sub-goals, and supporting tasks. The objective was to determine if there would be any impact on current TTPs or SOPs as a result of having these capabilities. This review did not generally change the nature of the TTP or derived measures.

SA Metric Development

After development of the general measures of performance and effectiveness using the GDKET, the MOUT ACTD program commissioned the U.S. Army Research Laboratory's Human Research and Engineering Directorate (HRED) Fort Benning Field Element to develop and validate an SA assessment metric for use in a MOUT ACTD field exercise as an objective measure of SA. The purpose of this measure was to evaluate the effect of the MOUT ACTD technology on the unit's SA. A unit equipped with current technology was used as the baseline in this assessment. Before development of the metric, several decisions had to be made. First, the type of assessment techniques to be used had to be selected. Second, the decision had to be made concerning the time of administration of the technique. Third, the type of exercise had to be decided.

Questionnaire Assessment of Knowledge Technique

Endsley (1995b) proposed that the ability to objectively measure SA is critical for progress and understanding in the field. She critiqued several measurement techniques that have been performed in the past to objectively measure SA. These include physiological techniques such as electroencephalographic measurements; performance measures used to infer SA (e.g., time to complete a scenario, loss exchange ratio, etc.); global measures of overall operator performance which give the end result of a long string of cognitive processes; subjective techniques such as self-rating and observer rating; and questionnaires about SA knowledge that evaluate against reality. She suggested that the questionnaire method provides an objective and direct assessment of SA.

Freeze-frame Technique

The questionnaire method can be administered during several different points in an exercise. It can be administered at the end of an exercise, during the conduct of an operator's simulated tasks, or using a freeze-frame technique. Endsley (1995b) found the freeze-frame technique to be more timely than the post test questionnaire and less disruptive than the on-line questionnaire. However, other authors have found fault with the freeze-frame technique. An often-stated criticism of the freeze-frame technique is that it is intrusive since it induces a temporary halt in the scenario (Sarter & Woods, 1991). However, in a simulation using fighter pilots, Endsley (1995b) found that the freeze-frame technique did not affect subjects' performance. She stated that the subjects' SA did not have a chance to decay during the freeze-frame before the SA simulation resumed. In other words, the SA was still intact and the freezing of the scenario did not have an adverse effect on the outcome of the scenario.

Free-play Exercises

Realistic aviation simulators have facilitated the study of SA in the aviation field. Pilots can be placed in the cockpit of simulators that are almost indistinguishable from the real thing. The pilots can then be presented with stimuli that are carefully controlled and processed. The response choices to these stimuli are few and the correct response choice is known when the simulation is developed. However, realistic simulations of the full range of infantry activities that allow assessment of SA have not been developed. The infantry environment is extremely dynamic and interactive, with many possible decision choices. An infantry simulation exercise cannot be limited to a single infantryman, because he operates as part of a unit or team and against other units or teams. Team SA requires a much more complex assessment than does combining the assessment of SA of individual team members. It requires assessment in its own right because it involves unique activities such as coordination and information sharing (Salas, Prince, Baker, & Shrestha, 1995). The decision choices in such an exercise are often numerous and they result in a decision choice matrix that is very complex. Also, outcomes are numerous and cannot be predicted before the exercise. For example, a scenario planner has no way of knowing how often an infantryman may shoot during an exercise and thus does not know ahead of time the correct answer to the SA question "how many rounds of ammunition do you have left?" A further complication of the ability to simulate the infantryman's environment is the fact that the infantryman is "moving, shooting, and communicating" while he attempts to maintain SA. He is not contained in an encapsulated environment such as a cockpit. Because of these factors, the only current way

to realistically replicate most of the critical aspects of the environment is in a free-play exercise.

Free-play exercises have been used for many years to train soldiers in combat skills. Units start training using free-play exercises at the squad level. They conduct squad exercises to practice their basic infantry skills (e.g., fire and maneuver, navigation, and weapon skills). Once these skills are perfected, squads come together under a platoon leader to test their individual training against an opposing force in a free-play exercise. The platoons then unite under a company commander to practice. The final evaluation of a commander is the evaluation of his unit during a force-on-force free-play exercise against a well-trained opposing force. The use of free play presents the unit commanders with an almost infinite number of decision possibilities and demonstrates the cause and effect outcome of the decisions.

Metric Development

Once the decisions were made to use questionnaire direct knowledge assessment, freeze-frame timing, and a free-play exercise, the HRED Field Element began work to develop the specific SA metric that would be used during the assessment. Development of the SA metric involved the following essential steps.

First, the developer acquired as much knowledge as possible of all facets of the scenarios planned. This included knowledge of the terrain on which the event took place; knowledge of the operations order that was provided to the unit to include objectives, constraints, and the time of day the operation would take place; and knowledge concerning the enemy opposing force such as the type of weapons they might carry and the number of enemy troops. The scenarios and threat used for the exercise were standard scenarios developed by the Army based upon a typical mission and threat. Scenarios that contain lots of action by either the enemy or the friendly forces are ideal for use in SA assessment.

Once knowledge of the scenario was gained, determination was made of how many freeze-frames were needed and where to place the freeze-frames in the scenario. Infantry SA is not static, but rather the result of ongoing processes within the unit. Therefore, a single measurement point was not adequate. SA assessment should be made over a series of important events while the unit is performing tasks (Salas et al., 1995). In a large exercise, the timing of each freeze-frame is critical. Endsley (1995b) found that subjects in her aviation study were able to provide information about their SA about a specific situation for as long as 6 minutes. Therefore, the freeze-frames were planned to take place no more than 6 minutes after critical situations occurred. Generally, the freeze-frame should provide as little disruption as possible. It was decided that the times just before naturally occurring breaks in a scenario were good locations for placement if these occurred within 6 minutes of critical SA elements.

Next, the developer identified critical SA elements that were contained within the specific event. This knowledge was acquired by holding an SME conference. A group of SMEs that were knowledgeable in the field of infantry and had experienced situations similar to those that would take place in the free-play exercise were best able to determine what the critical SA elements were for the specific exercise. These SMEs proposed SA elements that would be contained in the scenario and then prioritized them in terms of their importance within the scenario. Because it was desired that the amount of time during the freeze-frame would be kept to a minimum, only critical information was assessed. While many SA- critical elements are common across different types of scenarios, some are scenario dependent. Therefore, each set of SA queries was tailored to the specific type of scenario that was used. Not all critical SA questions were naturally occurring parts of a scenario. Sarter and Woods (1991) suggested that complex scenarios should include embedded events to elicit key situation assessment responses. The inclusion of several such events in the scenario allowed multiple opportunities for assessment. For example, an enemy operations order was left behind for the friendly soldiers to find. The transfer of this information to different levels of command was tracked through controlled queries at key points in the scenario.

The developer then determined the level of questions to present. Endsley (1987, 1988, 1995) defined SA as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future." Endsley's definition of SA discussed three levels: perception, comprehension, and projection. Perception (Level I SA) is the lowest level of SA because it involves only the conscious knowledge that something is present in the environment. Level 2 SA or comprehension is the synthesis of disjointed perceptions so that understanding of the significance of the perceptions is present. Projection (Level 3) SA is the ability to project future courses of action based upon the understanding gained from Level 2 SA. The critical SA elements identified for the scenario contained all three levels of SA, but the concentration was on Level I because more Level I critical elements naturally occur in a short time period.

Development of the specific SA questions to be administered to the soldiers was the fourth step. While many of these are natural extensions of the critical SA elements that were developed in Step 3, the explicit wording of the questions is critical to the assessment outcome. For example, when asking a platoon leader how many wounded he has, it is important to specify whether you mean in the headquarters' element or in the entire platoon. It is also informative to have the

respondent identify the source of his or her information for each SA question. The source question identifies from whom (i.e., platoon leader, squad leader, rifleman, etc.) or what (i.e., visual acquisition, auditory acquisition, over the radio, etc.) the information was gleaned. For example, if the purpose of the SA evaluation is to assess the contribution of a technology to the respondent's SA, then a source question can assist in the determination of whether the technology facilitated an increase in SA or whether it was another factor or another type of technology.

Once all the questions were refined, the next step was the definition of "ground truth." Ground truth is defined as the actual or "true" battlefield situation. It serves as the basis of comparison for what the subject perceives the situation to be. This is the most difficult and one of the most critical aspects of the development of a SA free-play experiment. Without an accurate definition of ground truth, an evaluation of SA cannot be made. In a free-play exercise, ground truth definition can come from a number of sources. If you have an instrumented facility, ground truth can come from position location devices or video cameras. Even if instrumentation is available, it is good to have a backup source of ground truth. Evaluator controllers are SMEs who accompany the unit members (friendly and enemy) during the exercise and are an excellent source of ground truth. Therefore, for each SA question, a matrix was developed to illustrate the source for ground truth. Sometimes more than one source was developed. For example, if the SA question asked a platoon leader how many enemy were on his objective, ground truth was obtained from video cameras on the objective, from position location devices on the enemy troops, and from evaluator controllers co-located with the enemy.

The final step was the SME ground truth questionnaire that was developed after the SA questionnaire via the ground truth source matrix. A different questionnaire was developed for the evaluator controllers at each location to gather information about what happened during each freeze-frame. For example, the questionnaire for the evaluator controller located with the opposing force addressed the questions in the SA questionnaire concerning the opposing forces. Questions covered such things as how many opposing force were killed during the preceding frame, where the opposing force was located during the proceeding frame, and whether any civilians were co-located with them. The questionnaire for the evaluator controller located with the platoon leader asked questions concerning his location during the previous frame and any orders he may have given during that time. The ground truth questions must be very specific and carefully worded just like the SA questions, or the responses will not always be useful. At least one ground truth question should be developed for every SA question identified in the ground truth matrix as being addressed by the questionnaire.

SA Data Analysis

After the exercise, the data gathered from the ground truth questionnaires and from the instrumentation was used to develop an answer sheet to use in scoring the SA questionnaires. Once the SA questions were scored, the percentage of correct answers was computed for each soldier (Marshak, Kuperman, Ramsey, & Wilson, 1987). Percentages were also computed by level of command (squad leader, platoon leader, and company commander) and by battlefield operating system question categories (i.e., maneuver, command and control, mobility and survivability, intelligence, combat service support, etc.). Chi-square tests were used to distinguish between the baseline and the MOUT ACTD technology SA levels.

Validation of the Metric

Drawing upon the writings of Schneider and Schmitt (1986), the military SMEs analyzed the content validity in conjunction with development of the GDKET measures of effectiveness and with the development of the SA free-play metric. A job analysis was performed by experts participating in the evaluation of content validity and domain sampling was used to represent the behaviors or knowledge skills and abilities found important for success. The free-play metric demonstrated both content and face validity.

Limitations

Several limitations were present during the SA experiments. Because of the nature of the free-play exercise, many uncontrolled variables may have affected the results obtained. These included leadership style, bad weather, late nights, long days, and their potential impact upon troop morale. Also, a free-play exercise involving teams (i.e., other squad members, other squads, and other platoons) does not control for the effect of the other individuals' expertise and motivation upon SA. Last, but not least, the exercises used during the experiments lasted for long periods of time (exceeding 1 hour). There was a desire on the part of the experiment directorate to disrupt the flow of the battle as little as possible. Therefore, only three freeze-frames were executed throughout the entire exercise. This resulted in many SA elements not being assessed.

Conclusion

In this article we sought to describe the development and application of methodologies used to develop operationally based SA measures and assess the contribution of MOUT ACTD technologies to the SA of individual soldiers and small units during free-play exercises. The methodologies discussed in this paper were successfully used to elicit field expedient measures of SA and to assess overall performance of individual and small units during experiments. The freeze frame methodology, in particular, was a valuable tool employed during several MOUT ACTD experiments. It demonstrated the ability to discriminate between baseline and technology conditions and the ability to track a learning curve over time. It also demonstrated both

content and face validity. Because the methodology was shown to have merit, it will be used in a large, company-sized, upcoming experiment that focuses exclusively on SA. This experiment will address TTPs for use primarily with intra-squad radios and the contribution of intra-squad radios to the SA of the squad.

References

Endsley, M. R. (1987). SAGAT: A methodology for the measurement of situation awareness (NOR DOC 87-83). Hawthorne, CA: Northrop Corp.

Endsley, M. R. (1988). Situation Awareness Global Assessment Technique (SAGAT). In Proceedings of the National Aerospace and Electronics Conference (pp. 789-795). New York: IEEE.

Endsley, M. R. (1995a). Toward a theory of situation awareness in dynamic systems. Human Factors, 37,32-64.

Endsley, M. R. (1995b) Measurement of situation awareness in dynamic systems. Human Factors, 37, 65-84.

Endsley, M. R. & Rogers, M. D. (1994), Situation awareness information requirements for en route air traffic control (DOT/FAA/AM-94/27) Washington, DC: Office of Aviation Medicine.

Marshak, W. P., Kuperman, G., Ramsey, E. G., & Wilson, D. (1987). Situational Awareness in Map Displays. In Proceedings of the Human Factors Society 31st Annual Meeting. (pp 533-535). Santa Monica, CA: Human Factors and Ergonomics Society.

Salas, E. Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation Awareness in Team Performance: Implications for Measurement and Training. Human Factors, 37, 123-126.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. International Journal of Aviation Psychology, 1, 45-57.

Schneider, B. & Schmitt, N. (1986). Staffing organizations. Glenview, IL.: Scott, Foresman.